Psychometric analysis of multiple-choice questions in an innovative curriculum in Kingdom of Saudi Arabia

Karim Eldin M. A. Salih[1,2], Abubakar Jibo[3], Masoud Ishaq[2], Sameer Khan[4], Osama A. Mohammed[5], Abdullah M. AL-Shahrani[3], Mohammed Abbas[1,2]

[1]Department of Pediatrics, [2]Medical Education, [3]Family and Community Medicine, [4]Physiology and [5]Pharmacology, College of Medicine, University of Bisha, Saudi Arabia

Abstract

Background and Aims: Worldwide, medical education and assessment of medical students are evolving. Psychometric analysis of the adopted assessment methods is thus, necessary for an efficient, reliable, valid and evidence based approach to the assessment of the students. The objective of this study was to determine the pattern of psychometric analysis of our courses conducted in the academic year 2018-2019, in an innovative curriculum. **Methods:** It was a cross-sectional-design study involving review of examination items over one academic session -2018/2019. All exam item analysis of courses completed within the three phases of the year were analyzed using SPSS V20 statistical software. **Results:** There were 24 courses conducted during the academic year 2018-2019, across the three academic phases. The total examination items were 1073 with 3219 distractors in one of four best option multiple choice questions (MCQs). The item analysis showed that the mean difficulty index (DIF I) was 79.1 ± 3.3. Items with good discrimination have a mean of 65 ± 11.2 and a distractor efficiency of 80.9%. Reliability Index (Kr20) across all exams in the three phases was 0.75. There was a significant difference within the examination items block (F = 12.31, F *critical* = 3.33, *P* < 0.05) across all the phases of the courses taken by the students. Similarly, significant differences existed among the three phases of the courses taken (F ratio = 12.44, F *critical* 4.10, *P* < 0.05). **Conclusion:** The psychometric analysis showed that the quality of examination questions was valid and reliable. Though differences were observed in items quality between different phases of study as well as within courses of study, it has generally remained consistent throughout the session. More efforts need to be channeled towards improving the quality in the future is recommended.

Keywords: Item analysis, courses, phases of study, Saudi Arabia, UBCOM

Introduction

Curriculums are guides used by teachers in schools to assist in the education of students. It contains objectives, activities units and suggested materials to enhance learning.[1] Curricular innovation is a managed process of development whose main products are teaching (and testing) materials, methodological skills, and pedagogical values perceived as new by potential adopters.[2] It is a willed intervention, which results in the development of ideas, practices, or beliefs, that are fundamentally new. In innovated, integrated curriculum, designing multiple-choice questions (MCQ) for assessments is a complicated and time-consuming process.[3] MCQs are the most commonly used tool for assessment of students in different courses offered at undergraduate and postgraduate levels, and capable of yielding examination items from the contents of the taught courses.[4] These items, when critically analyzed, provides feedback to both tutors and students on performance on each test item.

Address for correspondence: Dr. Karim Eldin M. A. Salih, Departments of Pediatrics and Medical Education, University of Bisha, College of Medicine, Saudi Arabia, P. O. Box 1290 Bisha 61922, Kingdom of Saudi Arabia. E.mail: Karimeldin_salih@hotmail.com

Access this article online

Quick Response Code:

Website:
www.jfmpc.com

DOI:
10.4103/jfmpc.jfmpc_358_20

Psychometric analysis of any test is defined sequences of events to collect data from a test to determine its quality.[3] One of the importance of item analysis is to know the reliability or consistency of the test administered.[3,5] This will ensure accountability to the community by providing competent graduates. The reliability of a test indicates its consistency, homogeneity and ultimately acceptability as a tool of measurement.[5] In item analysis, item difficulty and its ability to discriminate between students who knows and those who do not know determine the quality of the examination.[6,7] Providing a reliable test with reasonable difficulty will result in a type of assessment that can derive learning.[8] According to Ebel,[8] in a classical test theory item analysis, discrimination index (DI) of greater than 0.2 is acceptable. However, other workers suggested that any value above 0.15 is acceptable.[8,9] The difficulty index (DIF I) is determined by the number of the candidates who got the answer right over the total number of the students. A reasonable test should have difficulty index (DIF I) in the range of 50-80%.[10] Some authors consider DIF I above 80% as high implying that the questions are easy. On the other hand, low DIF I (less than 30%) means that the questions are difficult with a pressing need to improve the quality of the test item.[11-13] Discrimination index (DI), also called biserial point correlation (PBS), describes the ability of an item to distinguish between high and low scorers.[14] It ranges between -1.00 and + 1.00.It is expected that the high-performing students select the correct answer for each item more often than the low-performing students. If, however, the low performing students got a specific item correct more often than the high scorers, then that item has a negative DI (between -1.00 and 0.00).[15] The difficulty and discrimination indices are often reciprocally related. However, this may not always be true. Questions having high DI-value (more straightforward questions) tend to discriminate poorly; conversely, questions with a low DI-value (harder questions) are considered to be good discriminators.[16] Discrimination index of 0.40 and above is excellent, 0.30-0.39 is reasonably good, 0.20-0.29 is marginal items (i.e. subject to improvement), and 0.19 or less is poor items (i.e. to be rejected or improved by revision).[12,17]

A general indicator of test quality is the reliability estimate usually reported on the test analysis printout. Referred to as KR-20 or Coefficient Alpha, it reflects the extent to which the test would yield the same ranking of examinees if re-administered with no effect from the first administration.[18] Reliability (R) range from 0.7-1.0 is considered by many authors as excellent and acceptable.[13,19,20] A distracter efficiency (DE) is the list of distracters that distract and in an MCQ of three distractors it is 100%, 66%, 33% or 0%, if all the three distractors are distracting, tow distractors chosen ,one distractor chosen or all distractors not chosen. A functional distractor is distractors that has been attempted by at least 5% or more of the students.[9,21] According to Ebel and Downing, only 38% of distractors on the tests are eliminated because < 5% of students select them.[8] He reported that the percentage of items with three functioning distractors in most tests ranged from only 1.1 to 8.4% of all items.[22] The ultimate Goal and objective of each medical institutions should primarily directed to provision of evidence based care of the patient, however proper assessment by conducting high quality psychometric analysis of our assessment will ensure this. In fact gratifying advantage of multiple choice questions (MCQs), is it's ability to provide immediate feedback to all partners. This feedback will be maximally utilized if it combined with psychometric analysis. As it has been observed by many authorities psychometric analysis is an effective tools in in deciding what number of options, keeping questions in bank and comparing students achievement.[23] Amani *et. al.* (2020) used psychometric analysis to compare quality of MCQs designed by residents in radiology program and that design by teacher.[23] To ensure fair response from the examinees, many factors should be considered: on the top of these factors is the reliability of the questions and the test, the quality of the questions, the validity of the test and finally the language back ground of the examinee. Psychometric analysis will provide some answers to these questions by tell us how many option is needed for those who uses English as second language and can alert us to possible areas for improvement.[23] The rational of this study is to through some light across MCQs adopted for assessment as an effective tool, to draw the attention of teachers and administrations to the importance of unflawed MCQs and checking it's psychometric pattern.

The Objective

The objectives of our study is to determine the psychometric analysis of our courses conducted in the academic year 2018-2019, in an innovative curriculum.

Methods

Study design

The study is cross-sectional by design.

Site of the study

The study site was the College of Medicine, University of Bisha. The college was recently established to graduate competent doctors for the Kingdom.

Description of academic and assessment process

There are three phases integrated longitudinally and horizontally, with a total duration of six years study period. The curriculum adopts as methods of assessment, MCQs assessment, best of four, structured short answer questions, objective structured practical examination (OSPE), and objective structured clinical examination (OSCE). Quality chain for process of assessment is achieved through students' assessment committee (SAC), departmental meetings, and College Board. It involves review of exam items held within the 2018/2019 academic session at the college.

Data collection

Data was collected from the examination office for six weeks by five-trained research assistants.

Data analysis

The data on exam item analysis was generated by an optic reader machine used in marking the MCQ, the Apperson Datalink 3000 manufactured by Apperson.com, USA. The research assistants that participated are academic staff of the college-trained on data extraction and review of exam items analysis. Tool for data collection was a semi-structured questionnaire that was earlier validated by a pretest. Adjustments were made to the tool to capture all information required to address the specific objectives of the study. The Questionnaire captured essential items of the examinations held namely difficulty index, discrimination index, point biserial statistics, discriminator indices and efficiency as well as the test reliability (Kruder-Richardson-20). The reliability through Kuder-Richardson formula 20 (KR-20), Difficulty and the Discrimination Indexes, and the distractors functionality were considered for each question.[23,24] MCQ data was entered into Excel sheets and transferred to SPSS version 20 for analysis. Categorical variables were presented in the form of frequency and percentages. Tests of association were used to find relationships between variables of interest. The three phases were compared to detect any difference existing between and within the phases using a two- way analysis of variance (ANOVA). Similarly, differences in the questions within and between courses was carried out using ANOVA. F distribution statistics was used to determine variations between and within these two factors of interest, with F ratios and critical values determined. Significant differences are observed where the *P* values are less than 0.05. The ethical clearance was sought from the College Ethical Committee of the University of Bisha, College of Medicine to address the issues of concern.

Results

The total number of courses taken in the three phases of the session was 24, with first phase having 33.3% of the courses, second phase constituting 45.8% of the courses and the third phase having 20.8%. The total number of exam items was 1073 in these three phases. Of these, Phase one constituted 27.5% (*n* = 323), Phase two items were 45.2% (*n* = 530) and Phase three 18.8% (*n* = 220). Each question was made up of multiple-choice answer questions type A with three distractors and a key answer to the questions. The total number of distractors in 1073 MCQ questions was 3219 [Table 1]. Analysis of the exam items taken for the courses across the three phases of the session revealed that 11.4% of the exams questions during the session under review were tough (<30%). The mean difficulty across the three phases was 11.6 ± 1.8SD. Exams items that were very easy (ease index >85%) were16.8% of the questions with mean ease index score of 16.6 ± 5.1. The proportion of exams questions that were within acceptable DIF I (30-85%) was 71.9% with a mean of 71.9 ± 3.3 [Figure 1]. The discriminatory index (DI) showed that 24.2% of the questions poorly discriminates (DI <0.15) and have a mean of 25.1 ± 10.7SD. Similarly, the mean of the items with good discrimination was 65.2 ± 11.2 across all the three phases. The Point Biserial Statistics showed that 38.7% of the exam questions were of poor construction (P_{bsr} < 0.2) [Table 1]. Distractor indices analyzed showed 6.5% of the questions had three nonfunctioning distractors (3nFD), 16.5% had two nonfunctioning distractors (2nFD), and 33.7% had one nonfunctioning distractor (1nFD), in the session under review. The proportion of exam items with nonfunctioning distractors across all the three phases was 19.1%. The distractor efficiency (DE) observed within the 1073 items was 80.9% [Table 1]. Reliability index Kr20 across the three phases ranges from 0.5-0.8, and the mean KR 20 index for the exam items in the three phases is 0.754 [Table 1]. One-tenth (10%) of the questions in Phase 3 negatively discriminate students' scores. About one-third of the questions (28.6%) have zero discrimination as shown in Figure 2. The results of the ANOVA across the three phases are as shown in Table 2. F ratio for the rows (exam questions) with df = 2 was 12.32 which is higher than the F critical value of 3.23. A significant difference (*P* < 0.05) was seen within the problem questions across all phases of the courses taken by the students. Similarly, the F ratio for the columns (phases) was 12.44, which is higher than the F critical value (4.10). Thus, rejecting the null hypothesis and interpreted as having significant difference existing between the three phases of the courses (*P* < 0.05) [Table 3].

Table 1: Summary of all the phases combined																								
Course	Phase one								Phase two											Phase three					TOTAL
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
Ques	30	40	50	40	50	30	43	30	70	60	50	50	60	60	40	50	30	60	60	30	30	30	30	40	1073
Dif Q	2	5	4	12	9	3	6	3	9	6	5	8	12	5	3	4	1	1	9	0	1	3	5	6	122 (11.4%)
Eas Q	5	3	8	3	5	6	3	2	15	9	7	9	0	9	12	15	2	23	12	7	8	6	9	2	180 (16.8%)
Negd	3	2	1	5	1	0	0	0	3	3	3	0	7	0	1	5	1	1	9	0	1	1	5	5	58 (5.4%)
1nFD	16	14	12	10	17	15	10	11	29	16	19	19	13	22	13	19	10	15	26	11	15	7	7	16	362 (33.7%)
2nFD	7	2	4	7	6	2	3	7	17	11	14	9	4	5	8	11	5	21	7	6	5	5	10	5	181 (16.9%)
3nFD	3	2	3	2	2	5	1	1	3	5	1	3	0	6	5	10	0	13	1	2	1	1	0	0	70 (6.5%)
Rel.	0.6	0.7	0.8	0.8	0.8	0.7	0.9	0.8	0.9	0.9	0.8	0.9	0.9	0.8	0.8	0.7	0.8	0.8	0.7	0.8	0.8	0.6	0.5	0.7	

Key: Dif Q- Difficult questions, Eas Q- Easy questions, Negd -Negatively discriminating questions, 1nFD- 1 nonfunctioning distractor, 2nFD-- 2 nonfunctioning distractor,3nFD - 3 nonfunctioning distractor, Rel- Reliability Kr index, Dist- MCQ Distractors

Table 1: Shows that more than a third of the questions (33.7%) have one non-functional distractor, 16.9% have two non-functional distractors, and 6.5% of the total questions have three non-functional distractors. The reliability index Kr20 across the three phases ranges from 0.5-0.8.

Discussion

The students' performance in the MCQ of these 24 courses was used to determine the difficulty index, discrimination index and non-functional distractors or study evaluates how the MCQ differentiate between student's performance, within the test items in each of the courses and across the three phases of study. In this study, the DIF I of the items was 71.9% similar to what was reported by other studies,[3,10,25,26] where 61-80% of the items are within acceptable range. However, our study revealed the difficulty and ease indices in the three phases are lower than those reported by these studies. The discriminatory index (DI) shows up to a quarter of our assessment's items (25.1%) poorly discriminates between good and poor students (DI <0.15). Some authors[9,10] reported range of 14-17% of exam questions were poorly discriminating compared to about a quarter (25.1%) seen in our study. This difference could be partly due to the variation in the cut-off score adopted for the studies and partly ascribed to the different tools used for the study. About sixty-five per cent

of the items showed excellent discrimination (>0.2). The DI index is similar to that reported by Rao *et al*.[27], where 60% of the items were good discriminators. Other authors[28,29] have reported that DI of 0.2 is acceptable and would discriminate between weak and good students. We observed a distractor efficiency of 80.9% and non-functioning distractors accounting for 19.1% of questions across all phases. Two-third of the questions have one non-functional distractors, 16.9% have two and less than a tenth (6.5%) have three non-functional distractors. Tarrant in Hong Kong[21] reported similar to that where 13.8% of the total items he tested had only three functioning distractors, where 70% of the items had one or two functional distractors.[15] Other authors documented more than 66% of items showed non-functional dissectors [Table 3]. There was significant difference in difficulty factor, discrimination indices, and reliability and distractors functionality between different phases. The probable reasons for these results could be meticulous internal regulations adopted by the SAC, departments and course coordinators before and after conduction of the examinations. Weekly feedback to tutors

Table 2: All phases combined

Problem Questions	Phase 1	Phase 2	Phase 3	Row Total	Row average
No of diff Q	44	54	24	122	40.67
No of easy Q	35	101	44	180	60
No of neg discr Q	12	24	21	57	19
1 NFDs	105	175	82	362	120.67
2 NFDs	38	105	38	181	60.33
3 NFDs	19	46	5	70	23.33
Total	253	505	214	972	
Column average	42.17	84.17	35.67		

Key: No of Diff Q- Difficult questions, No of Easy Q- Easy questions, No of Neg -Negatively discriminating questions, 1nFD- 1 nonfunctioning distractor, 2nFD--2 nonfunctioning distractor,3nFD - 3 nonfunctioning distractor, Rel- Reliability Kr index, Dist- MCQ Distractors

A two-way analysis of variance was done to determine the differences between the three phases of study as well as differences in the problem questions answered during the three phases.
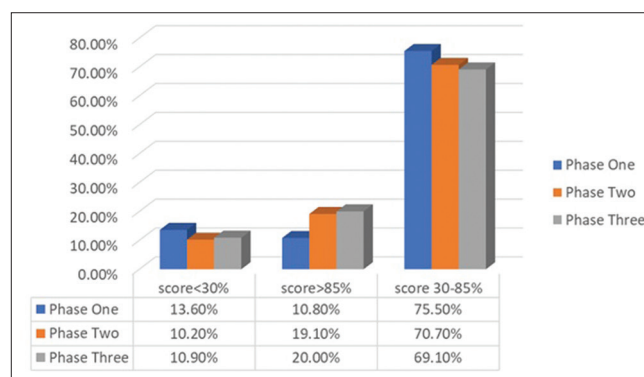

Figure 1: The figure shows examination difficulty index observed across the three phases of the courses taken. The mean DIF I was 71.8% across the three phases

Table 3: Anova: Two-factor without replication

SUMMARY	Count	Sum	Average	Variance			
Diff Q	3	122	40.66667	233.3333			
easy	3	180	60	1281			
Neg DSC	3	57	19	39			
1nFD	3	362	120.6667	2346.333			
2nFD	3	181	60.33333	1496.333			
3nFD	3	70	23.33333	434.3333			
P1	6	253	42.16667	1093.367			
P2	6	505	84.16667	2990.967			
P3	6	214	35.66667	702.6667			
ANOVA							
Source of Variation	SS	df	MS	F	P	F crit	
Rows	20591.33333	5	4118.267	12.31662	0.000518	3.325835	
Columns	8317	2	4158.5	12.43695	0.001939	4.102821	
Error	3343.666667	10	334.3667				
Total	32252	17					

The ANOVA (two way) results shows that the F ratio for the rows (exam questions) with df=2was observed to be 12.32 is higher than the F crit value of 3.23. It means a significant difference (*P* < 0.05) exist within the problem questions block across all the phases of the courses taken by the students. Similarly, the F ratio for the columns (phases) is 12.44, which is higher than the F critical value (4.10). Thus, rejecting the null hypothesis and interpreted as significant difference existing between the exam items in the three phases of the courses (*P* < 0.05).

No Q- number of questions. SS= sum square. Diff Q-Difficult questions. df =degree of freedom. Easy Q- easy questions. MS=Mean square. Neg DSCR- Negative discrimination question. F= F ratio. 81NFD- 1 non-functional distractor. F crit=F critical value. 2NFD- 2 non-functional distractor. 3NFD-- 1 non-functional distractor. P1-Phase one. P2-Phase two. P3- Phase three
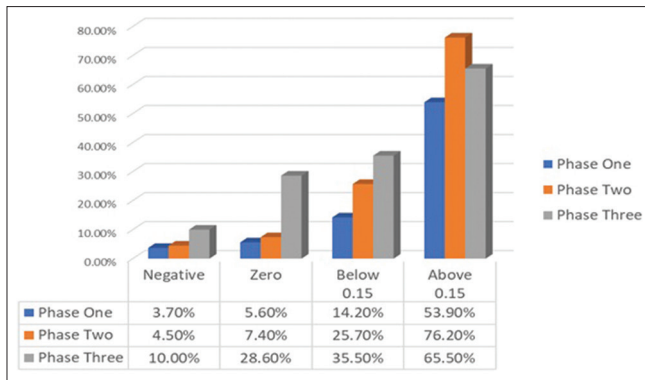
	Negative	Zero	Below 0.15	Above 0.15
■ Phase One	3.70%	5.60%	14.20%	53.90%
■ Phase Two	4.50%	7.40%	25.70%	76.20%
■ Phase Three	10.00%	28.60%	35.50%	65.50%

Figure 2: One-tenth (10%) of the questions in phase three negatively discriminate students' scores. about one-third of the questions (28.6%) have zero discrimination as observed

is known to be essential in training,[30] this is made possible through panel discussions and other activities in our innovative curriculum. It improves the quality of training and understanding on the part of the students. It is obvious from this study and other studies adherence to restrict MCQs guidelines in general, beside application of cover hand test i.e. to anticipate the true answer even without looking at the options can provide better psychometric analysis regardless of option numbers.[24] To the knowledge of the authors unfocussed questions in general and which of the fallowing types of questions in particular will affect the quality of psychometric analysis.

Conclusion

Psychometric analysis of exam items showed that the quality of examination questions was valid and reliable. Variations in items quality have been observed between different phases of study as well as within courses of study that the quality of the exam items has generally remained consistent throughout the session.

Recommendation

Psychometric analysis is urgently needed so as to determine area of improvement and to build reliable MCQs bank.

Limitation

The data was collected from one institute during one academic year and numbers of students were less.

Strength

The work addresses the determination of the MCQs questions in an innovative curriculum.

Acknowledgement

Financial support and sponsorship

Conflicts of interest

There are no conflicts of interest.

References

1. Sherratt Patrice. Principles of Curriculum Development, Implementation and Evaluation. Available from: https://www.academia.edu › Curriculum Innovations. [Last accessed on 2019 Oct 12].

2. Barriga FD, Barrón C. Innovation in education and curriculum reforms in Mexico. Transnatl Curriculum Inquiry 2012;9. Available from: http://www. nitinat.library.ubc.ca/ojs/index.php/tci. [Last accessed on 2019 Oct 12].

3. Mitra NK, Nagaraja HS, Ponnudurai G, Judson JP. The levels of difficulty and discrimination indices in type A multiple-choice questions of pre-clinical semester one multidisciplinary summative tests. ASME 2009;3:2-7.

4. Afraa M, Samir S, Ammar A. Distractor analysis of multiple-choice questions: A descriptive study of physiology examinations at the Faculty of Medicine, University of Khartoum. Khartoum Med J 2018;11:1444-53.

5. Zubairi AM, Kassim NL. Classical and Rasch analyses of dichotomously scored reading comprehension test items. Malays J ELT Res 2016;2:20.

6. Warburton WI, Conole GC. Key findings from recent literature on computer-aided assessment. In ALT-C 2003. Sheffield; 2002.

7. McAlpine M, Hesketh I. Multiple response questions- allowing for chance in authentic assessments. In: Christie J, editor. 7th International CAA Conference. Loughborough: Loughborough University; 2003

8. Ebel RL. Essentials of Educational Measurement. 1st ed. New Jersey: Prentice-Hall; 1972.

9. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality multiple-choice questions (MCQs) from an assessment of medical students of Ahmedabad, Gujarat. Indian J Community Med 2014;39:17-20.

10. Elfaki OA, Bahamdan KA, Al-Humayed S. Evaluating the quality of multiple-choice questions used for final exams at the Department of Internal Medicine, College of Medicine, King Khalid University. Sudan Med Monitor 2015;10:123.

11. Dixon R. Evaluating and improving multiple-choice papers: True-false questions in public health medicine. Med Educ 1994;28:400-8.

12. Rahim AFbA. What Those Number Mean? 1st ed. Kubang Kerian: KKMED; 2010. Available from: http://www.medic.usm.my/dme/images/stories/staff/KKMED/2010/item_analysis_guide.pdf. [Last accessed on 2019 Sep 9].

13. Taib F, Yusoff MS. Difficulty index, discrimination index, sensitivity and specificity of long case and multiple-choice questions to predict medical students' examination performance. J Taibah Univ Med Sci 2014;9:110-4.

14. Fowell SL, Southgate LJ, Bligh JG. Evaluating assessment: The missing link? Med Educ 1999;33:276-81.

15. Mozaffer RH, Farhan J. Analysis of one-best MCQs: The difficulty index, discrimination index and distractor efficiency. J Pak Med Assoc 2012;62:142-7.

16. Carroll RG. Evaluation of vignette-type examination items

for testing medical physiology. Am J Physiol 1993;264:S11-5.

17. Ebel RL, Frisbie DA. Essentials of Educational Measurement. 5th ed. Englewood Cliffs, New Jersey: Prentice-Hall Inc; 1991.

18. Kehoe J. Basic item analysis for multiple-choice tests. Practical Assessment, Research and Evaluation 1995;4. Available from: http://PAREonline.net/getvn.asp?v=4 and n=10. [Last accessed on 2019 Oct 9].

19. Rahma NA, Shamad MM, Idris ME, Elfaki OA, Elfakey WE, Salih KM. Comparison in the quality of distractors in three and four options type of multiple-choice questions. Adv Med Educ Pract 2017;8:287-91.

20. El-Uri FI, Malas N. Analysis of use of a single best answer format in an undergraduate medical examination. Qatar Med J 2013;2013:3-6.

21. Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. BMC Med Educ 2009;9:1-8.

22. Haladyna TM, Downing SM. How many options is enough for a multiple-choice test item? Educ Psychol Meas 1993;53:999-1010.

23. Amini N, Michoux N, Warnier L, Malcourant E, Coche E, Berg BV. Inclusion of MCQs written by radiology residents in their annual evaluation: Innovative method to enhance resident's empowerment? Insights Imaging 2020;11:1-8.

24. Tweed M. Adding to the debate on the numbers of options for MCQs: The case for not being limited to MCQs with three, four or five options. BMC Med Educ 2019;19:1-4.

25. Karelia BN, Pillai A, Vegada BN. The levels of difficulty and discrimination indices and relationship between them in four-response type multiple choice questions of pharmacology summative tests of year II M.B.B.S students. IeJSME 2013;7:41-6.

26. Patel KA, Mahajan NR. Itemized analysis of questions of multiple choice question (MCQ) exam. Int J Sci Res 2013;2:279-80.

27. Rao C, Kishan Prasad HL, Sajitha K, Permi H, Shetty J. Item analysis of multiple-choice questions: Assessing an assessment tool in medical students. Int J Educ Psychol Res 2016;2:201-4.

28. Brown FG. Principles of Educational and Psychological Testing. 3rd ed. New York: Holt, Rinehart and Winston; 1983.

29. Crocker L, Algina J. Introduction to Classical and Modern Test Theory. New York: Holt, Rinehart and Winston; 1986.

30. Salih KMA, Al-Shahrani AM, Eljac IA, Abbas M. Perception of faculty members of regional medical school toward faculty development program. Sudan J Med Sci 2019;14:65-77.